FINAL REPORT

Tools for Risk-Based UXO Remediation

SERDP Project MR-1718

JANUARY 2014

Dean Keiswetter
Tom Bell
Leidos Holdings Corporation

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
JAN 2014	**N/A**	**-**

4. TITLE AND SUBTITLE
Tools for Risk-Based UXO Remediation

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Leidos Holdings Corporation

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The primary purpose of the process described and illustrated in this SEED study is to confirm that the detection and classification decisions made prior to and during a remedial effort remain reasonable once ground truth information is garnered and taken into account. Specific objectives were to investigate (i) a practical method for assessing target detections and (ii) procedures for determining the likelihood of false negatives given some amount of ground truth information.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**SAR**	**31**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Abstract

Objectives: The primary purpose of the process described and illustrated in this SEED study is to confirm that the detection and classification decisions made prior to and during a remedial effort remain reasonable once ground truth information is garnered and taken into account. Specific objectives were to investigate (i) a practical method for assessing target detections and (ii) procedures for determining the likelihood of false negatives given some amount of ground truth information.

Technical Approach: The problem of UXO residual risk is naturally divided into the issues of target detection and, once detected, classification. With regards to the assessment of missed UXO detections, we present a practical and straight forward tool that involves inserting synthetic signatures, appropriately chosen for the UXO types actually recovered, into the reconnaissance survey data, processing using standard detection schemes and thresholds, and evaluating the results in the light of ground truth information. This approach is attractive because it naturally incorporates site-specific realistic noise levels and sampling schemes.

The second issue, that of classification failures, results when TOI targets ”look” more like what they are not than what they are, and are classified accordingly. In fact, because the decision metric is based on a library of expected TOI signatures, the target does not look “enough like” a TOI to be considered TOI. To investigate the possibility of classification failures, we (i) performed a probabilistic risk assessment using polarizabilities and ground truth information from Camp San Luis Obispo, Camp Butner, and Camp Beale, (ii) conducted a cluster analysis to look for groups of targets with similar signatures, and (iii) calculated a decision metric based on clutter signatures and compared it with the decision metric based on the library of expected TOI signatures.

Results: Based on a comparison of the actual depth distribution of the UXO recovered at San Luis Obispo and results of the synthetic seed study, we conclude that all of the UXO, at least with regard to the four included in the study, were successfully detected. The deepest target that we know of, based on the available dig results, was an 81mm at a depth of 0.52m. The synthetic seed exercise suggests, however, that 100% of the 81mm mortar seeds would have been detected, using the same detection scheme, for burial depths of up to 0.77m. Thus, the detection process applied to ESTCP’s Classification Study at San Luis Obispo, CA, was validated, given all available ground truth information resulting from excavations.

Using data and ground truth information from Pole Mountain, we found that roughly half of the clutter would be dug at a threshold corresponding to a 1% residual risk of misclassified TOI, if extreme value theory is used to estimate the likelihood that one or more TOI may be misclassified. This application of statistics appears misguided. The mathematical formulations are correct but the logic is not. The outliers that dominate the risk calculations do not reflect

random sample variations in the decision metric, but rather are due to inversion failures or other pathological events. To pursue the nature of the outliers, we used a cluster analyses and the calculation of a decision metric based on clutter signatures to show that, with the exception of a single TOI at Camp Beale that doesn't match anything (a singleton), the TOI versus non-TOI classification decisions were reasonable.

Benefits: Although assumptions can be made based on land use, expected munitions, expected clutter, and site conditions prior to conducting a geophysical survey, significantly more information is available after the geophysical survey and characterization is complete. Information about the site is learned during all phases of the effort; from the reconnaissance survey, to the anomaly characterizations, through the classification, and finally, when ground truth information is gathered. This additional information can and should be exploited, using the processes and methods presented in this report, to re-evaluate the detection thresholds and classification decisions prior to site closure. By adopting such procedures, stakeholder confidence and acceptance can be increased.

Table of Contents

List of Figures

List of Acronyms

EMI	Electromagnetic Induction
ESTCP	Environmental Security Technology Certification Program
FOM	Figure of Merit
MM	MetalMapper
MR	Munitions Response
Rx	Receive
SERDP	Strategic Environmental Research and Development Program
TEM	Transient Electromagnetic
TEMTADS	Transient Electromagnetic Towed Array Discrimination System
TOI	Target of Interest
Tx	Transmit
UXO	Unexploded Ordnance

Keywords

Unexploded Ordnance

Detection

Classification

Residual Risk

Electromagnetic Induction

Polarizability

Acknowledgements

We gratefully acknowledge financial support from the Strategic Environmental Research and Development Program for this project.

Objective

The technical objective of this SEED project was to investigate a practical method that facilitates the assessment of residual UXO risk, expressed in terms of probability that one or more targets of interest remain onsite after a geophysical survey and cleanup operation is complete. Within this context, we investigated a practical method for assessing target detections and devised procedures for determining the likelihood of misclassifying a UXO as clutter.

Background

New technologies for identifying and classifying UXO from non-UXO have been developed during the recent past with SERDP and/or ESTCP funding. Successful classification results have been realized by multiple firms, using a variety of advanced electromagnetic induction sensors, at multiple live site demonstrations. As a result of the success, these hardware and software technologies are also being introduced and incorporated into active site cleanup projects because they offer the potential to substantially reduce overall remediation costs. The full impact of UXO classification will be realized, however, only if the wider community of stakeholders, regulators, and site managers accepts their usage.

In order for human health hazard of munitions to be successfully mitigated, the munitions present at a site must be detected, then classified correctly, and disposed of properly. Each of these three tasks are made increasingly more difficult if a wide variety of munitions types are present, it the site is highly cluttered with non-UXO metal or fragments, or if the geologic background response is strong. Although assumptions can be made based on prior land use, expected munitions, and site conditions prior to conducting a remediation, significantly more information is available once the area is surveyed, characterized, and ground truth information is garnered. This additional information can and should be used prior to site closure to re-evaluate the detection thresholds and classification decisions. The detection thresholds and classification decisions may change, for example, if one or more unexpected munitions types are recovered during the course of remedial investigation.

To encourage stakeholder acceptance, we believe that a principled assessment of residual UXO risk, defined as the probability that one or more buried UXO remain onsite after a geophysical survey and cleanup operation is complete, is required. Our approach for assessing residual UXO was to explicitly address two essential and fundamental questions during site closure discussions; namely, if given all available geophysical and ground truth data collected at the site in question:

(i) Is it reasonable to believe that all UXO were detected?
(ii) Does it appear that any potential UXO were misclassified as clutter?

We use data from ESTCP Live Site Demonstration sites to illustrate the concepts and methods.

Missed Detections

The primary goal of the process described and illustrated in this report is to confirm that the detection and classification decisions made prior to and during a remedial effort remain reasonable once ground truth information is garnered and taken into account.

In the following discussion, we assume that we have access to all geophysical data, both reconnaissance survey and classification, and all ground truth information, which includes recovered munitions types and burial depths, as well as all non-UXO discoveries. Given this scenario, our approach for assessing missed detections includes the following steps:

- First, insert synthetic signatures, which correspond to the expected response of the recovered munitions, directly into the geophysical survey data.
- Then, process using standard detection procedures.
- Finally, evaluate detection statistics as a function of depth and munitions types, etc.

This approach naturally incorporates realistic noise levels and scenarios found at the area of concern and imposes realistic sampling scenarios.

Camp San Luis Obispo

Data from the former Camp San Luis Obispo, California was selected for this study. The site spans a hillside that is a historical mortar target and has limited vegetation and geologic interference, and a variety of munitions types. At this site, there were four targets of interest known prior to the study: 60-mm, 81-mm, and 4.2-in mortars and 2.36-in rockets (Figure 1).

Figure 1. Photographs of the seeded targets at Camp San Luis Obispo: top left – 2.36in rocket, top-right -- 4.1in mortar, bottom-left—81mm mortar, and bottom-right –60mm mortar.

A photograph showing the local setting of the San Luis Obispo Classification Study site is shown in Figure 2. Historic evidence and recent field observations document intact munitions and debris in this area. The average anomaly density was expected to be 100-200 anomalies per acre, which is low enough to allow for classification of individual items.

Figure 2. Photograph of the study area within Former Camp San Luis Obispo, California.

During the initial ESTCP Demonstration study, a team from the ESTCP program office produced a list of anomaly detections. This list was then passed to all of the classification demonstrators to attempt classification. Overall, the results of the classification efforts were very good. Additional details and classification results can be found in [1].

Here, we are interested in quantifying the likelihood that targets of interest were not identified during the detection phase. The evaluation process that we followed included these steps:

1. Run an automatic target picker on EM61-MK2 data, which was used for the detection part of this study, to determine appropriate threshold to detect original anomalies. Note that this step was required primarily because we did not perform the target detections and did not have access to their detection parameters.
2. Mask (remove) data around each of the original anomalies. The reason for removing the original detections is that we were interested in missed detections – and those anomalies nearby or on top of already detected anomalies were, by definition, not missed.
3. Randomly select seed locations from the remaining background data to plant model responses of the UXO type found at the site. For these example data, we were able to identify ~4,100 seed locations.
4. Using the site data calculate average polarizations for each UXO type.
5. At each seed location, add synthetic signatures, e.g., the expected response associated with each individual munitions type, to the measured data. Note that the measured data at each seed location is solely comprised of natural background "noise". For this study, we constrained the burial depth to be between 8 and 15 times the ordnance diameter.
6. Run automatic target picker on seeded data using different detection thresholds.

7. Calculate probability of detection statistics by creating bins which correspond to the individual gradations from 8-15 times ordnance diameter.

These steps are illustrated in the subsequent text and figures.

Figure 3 presents false-color images from the acquired EM61 data at San Luis Obispo, CA. The nominal line spacing for the EM61 data was 0.5 meter. Approximately 1500 targets were selected for intrusive actions during the study. Ground truth information regarding the nature of the source item at each of the 1500 detections was collected during the original study. Our task is to determine if these original detections are still reasonable given ground truth information regarding recovered munitions type and burial depths.

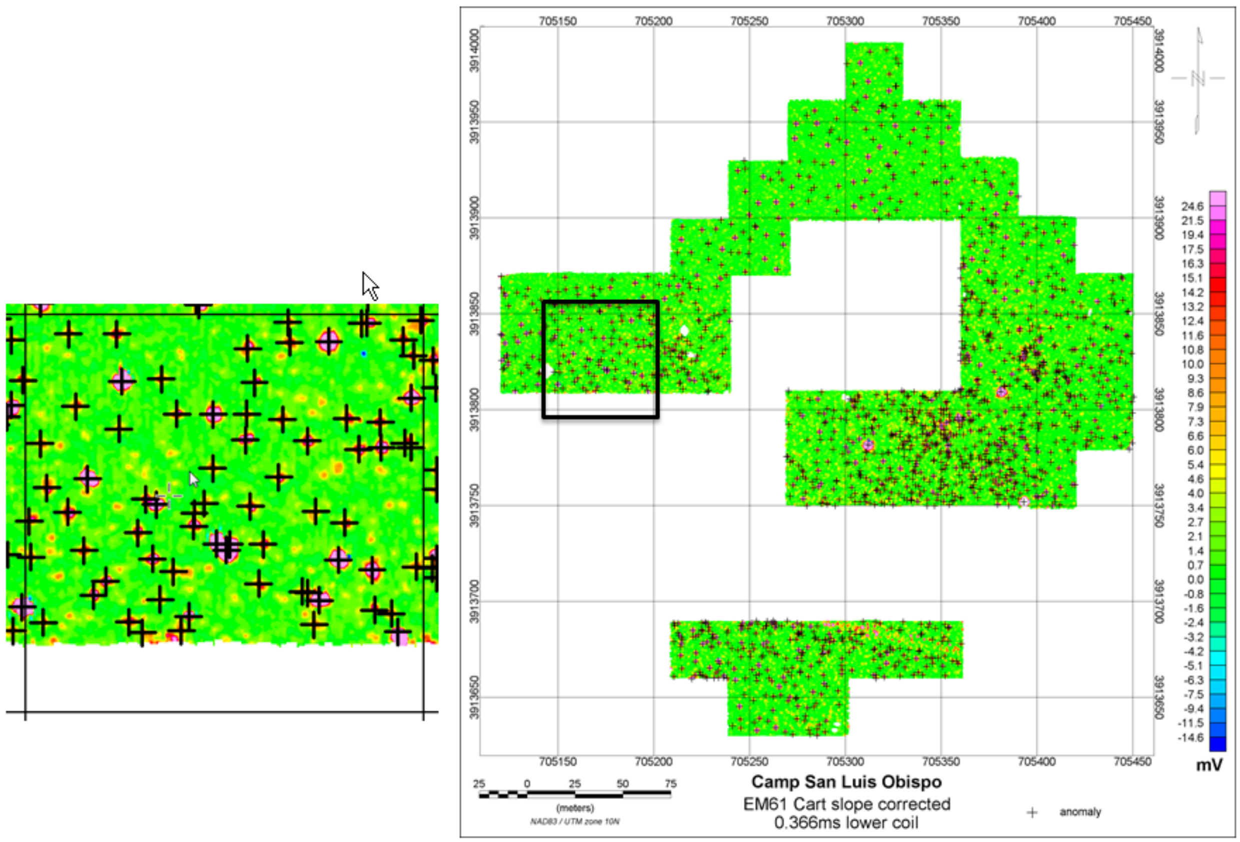

Figure 3. EM61 data from Camp San Luis Obispo overlain by target declarations as picked by the ESTCP program office for the initial demonstration study.

Figure 4 presents EM61 MKII cart survey data with original anomalies and automatic target picker declarations using 6mV threshold.

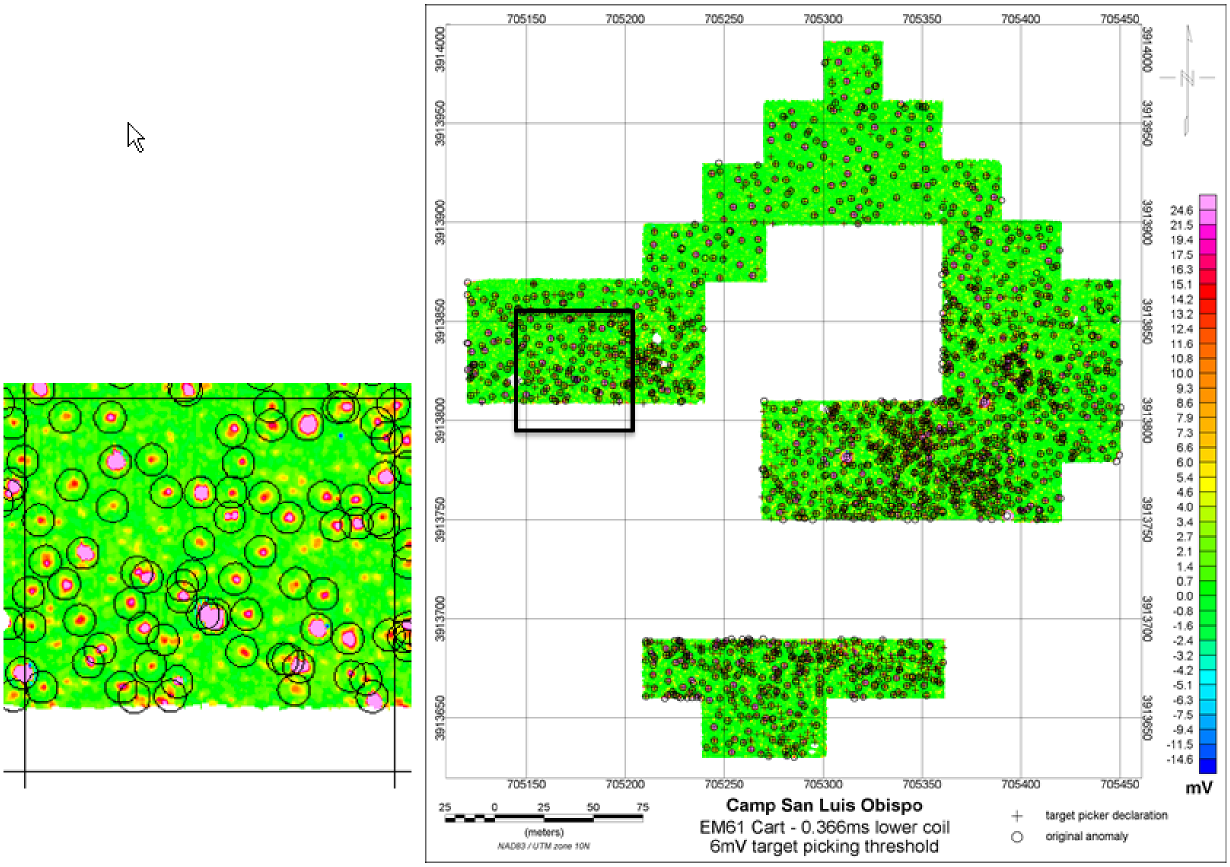

Figure 4. EM61 data from Camp San Luis Obispo overlain by target declarations made by an automatic target picker using a 6mV threshold.

The final step before seeding the site with synthetic targets is to remove data from each original target area (Figure 5). Once the anomalous responses were removed, the remaining areas were considered to represent natural background levels and fluctuations.

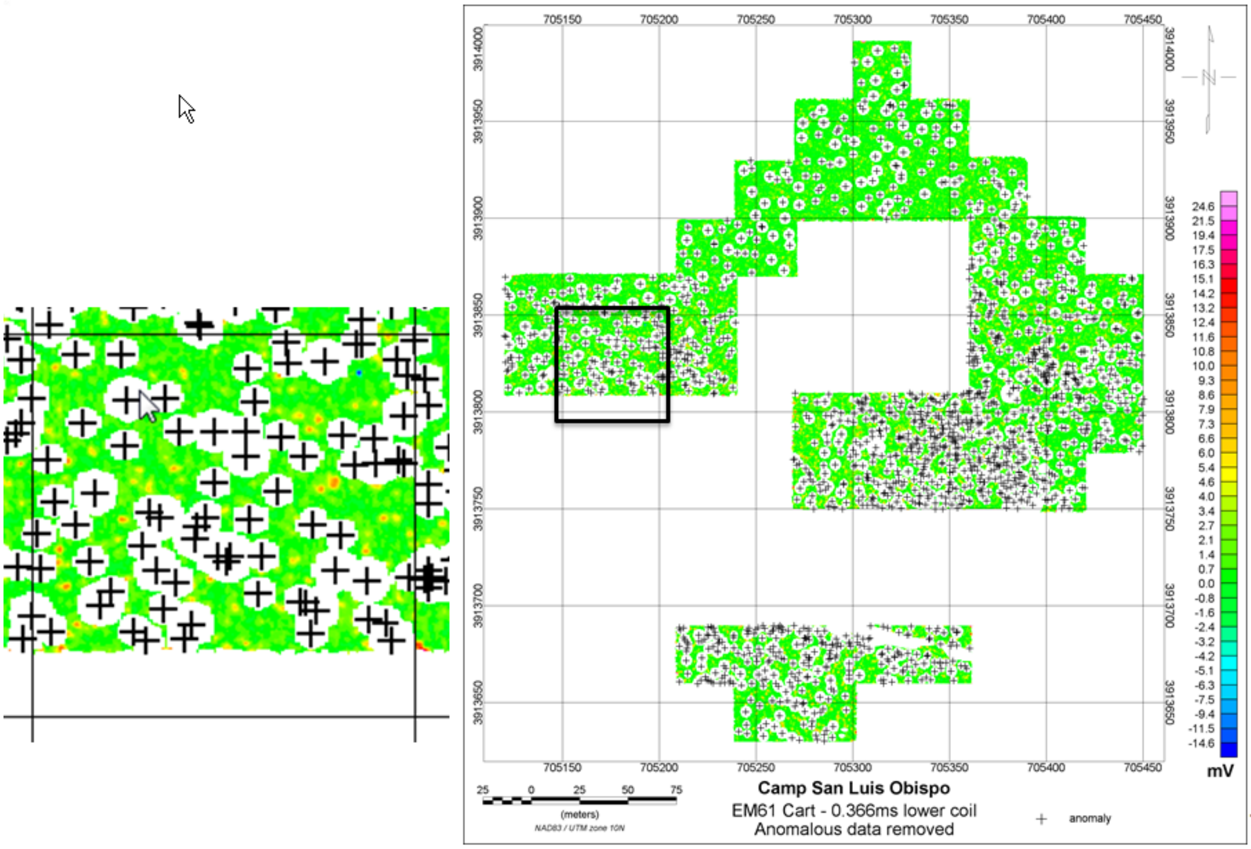

Figure 5. EM61 data after removing data associated with the original targets.

Figure 6 shows the EM61 MKII cart survey data with over 4100 randomly selected seed locations in background areas.

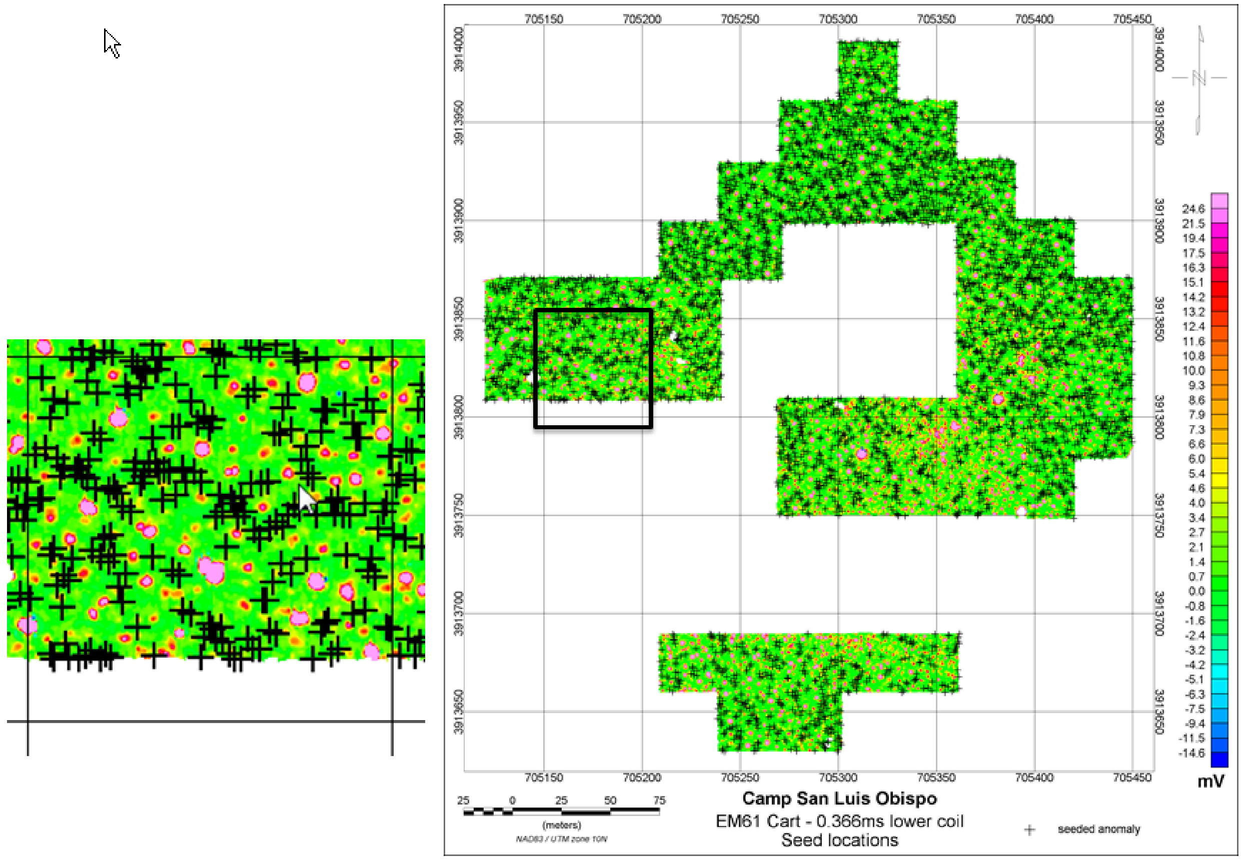

Figure 6. EM61 data from Camp San Luis Obispo overlain by locations for seeded targets.

For each seed location the response of each UXO type was added to the background (Figure 7). All target parameters for a UXO type were selected randomly except for depth which was constrained to be from 8 to 15 times the ordnance diameter.

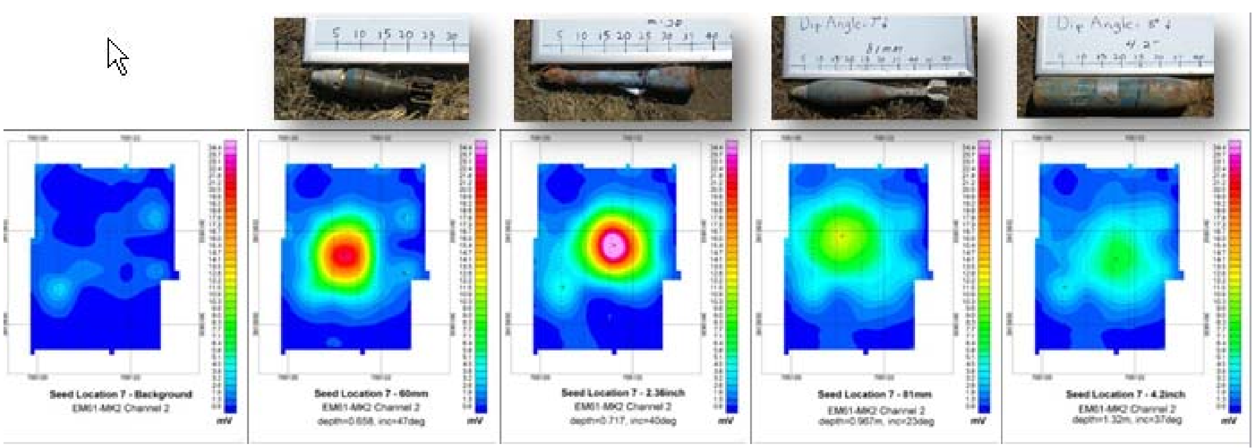

Figure 7. Decision metric distributions from recent ESTCP classification demonstrations.

Finally, we used an automatic target picker to select targets at thresholds of 2-, 4-, 6-, and 8mV. We accumulated the target declarations and calculated the probability of detection for each UXO type at the four thresholds. The depths correspond to bins ranging from 9 to 15 times the ordnance diameter.

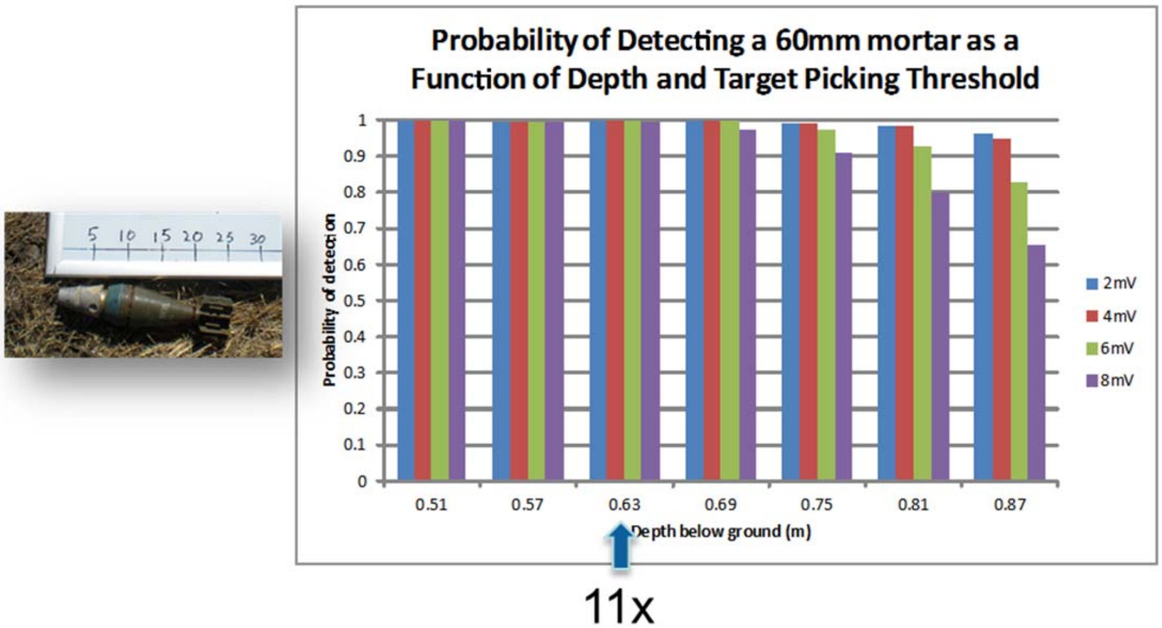

Figure 8. Probability of detection for 60mm mortar as a function of depth and target picking threshold for the San Luis Obispo data set.

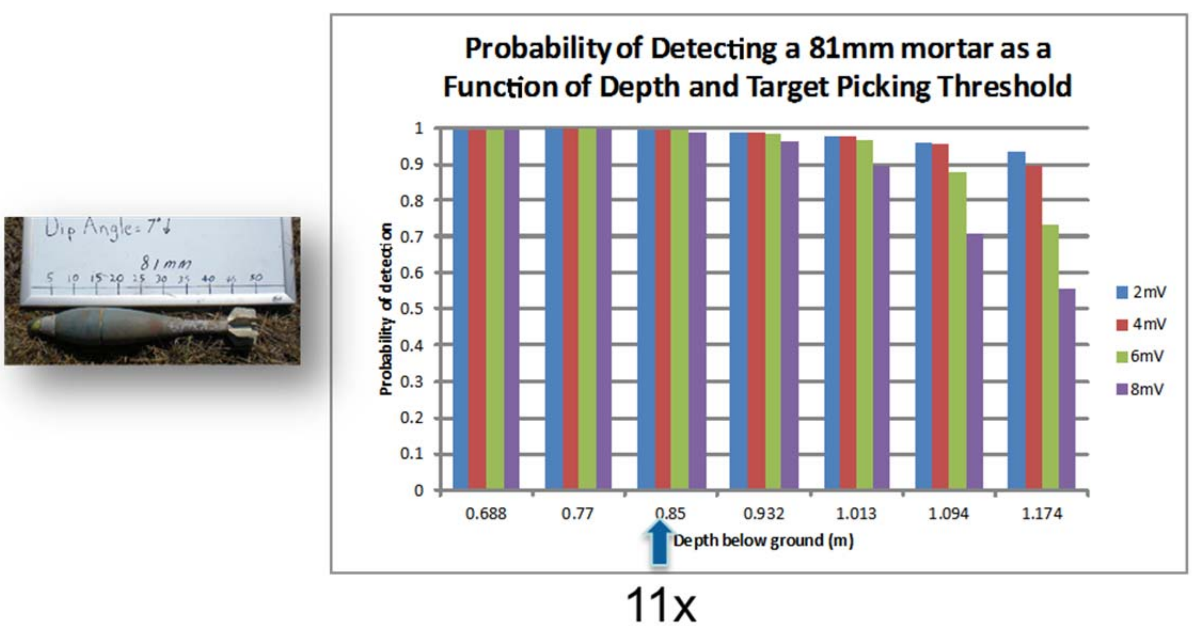

Figure 9. Probability of detection for an 81mm mortar as a function of depth and target picking threshold for the San Luis Obispo data set.

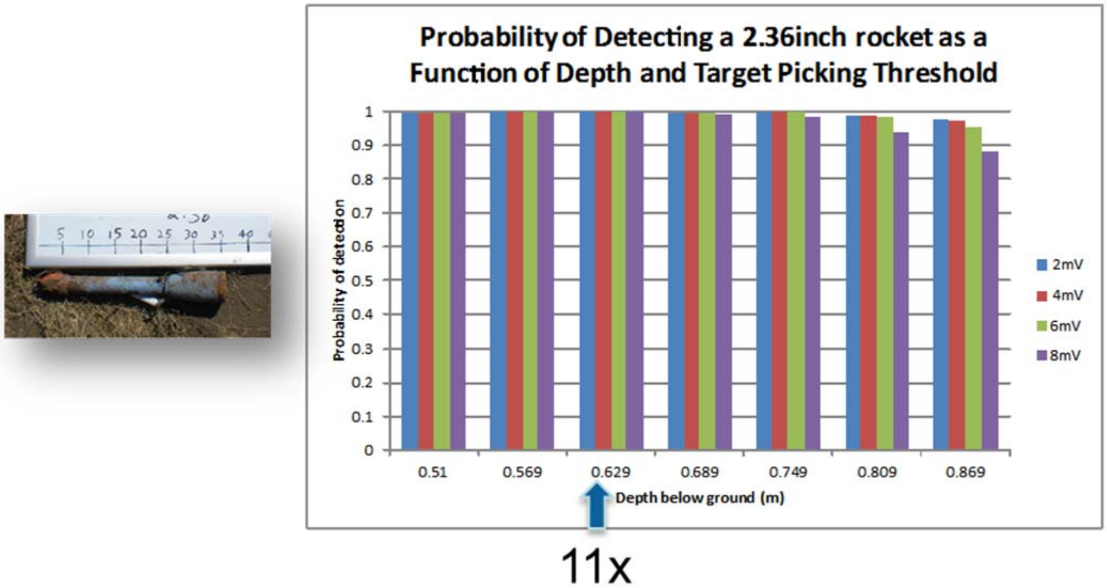

Figure 10. Probability of detection for 2.36in rocket as a function of depth and target picking threshold for the San Luis Obispo data set.

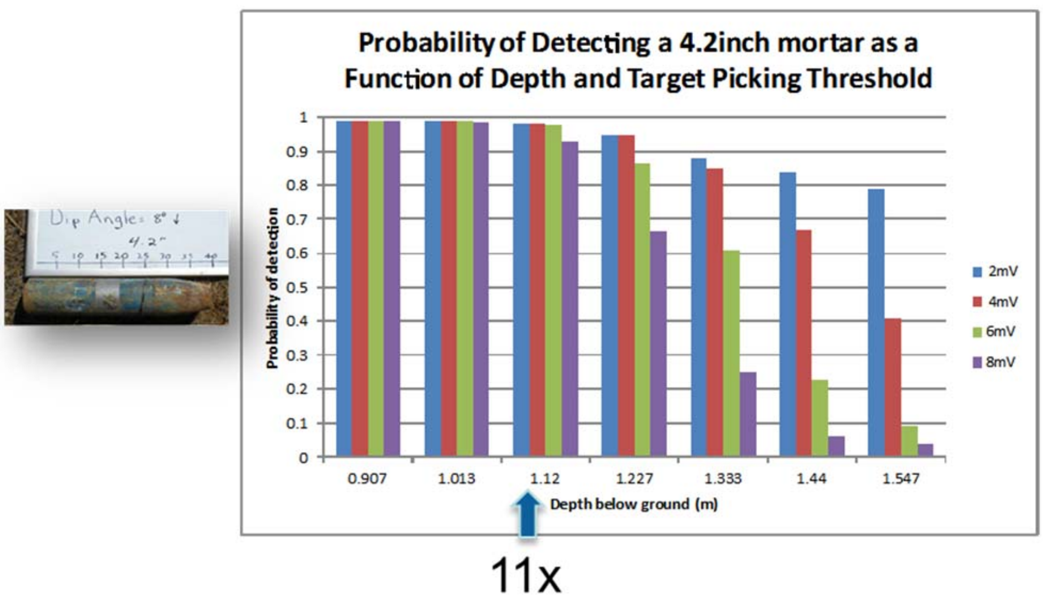

Figure 11. Probability of detection for 4.2in mortar as a function of depth and target picking threshold for the San Luis Obispo data set.

As observed in these results, there are occasional missed declarations for the deeper burial depths. The misses are most notable for targets at depths of 13x the diameter, but occasionally for shallower depths. We investigated these misses and found that the missed detections at

shallow depths were due to seeded targets located near data gaps (Figure 12). The missed target declarations at deeper depths were due to low signal to noise, which is expected.

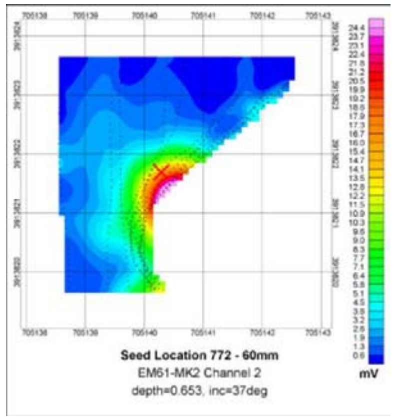

Figure 12. Example of a shallow synthetic seed that was not detected due to partial coverage (sampling). The red 'X' is the actual seed location.

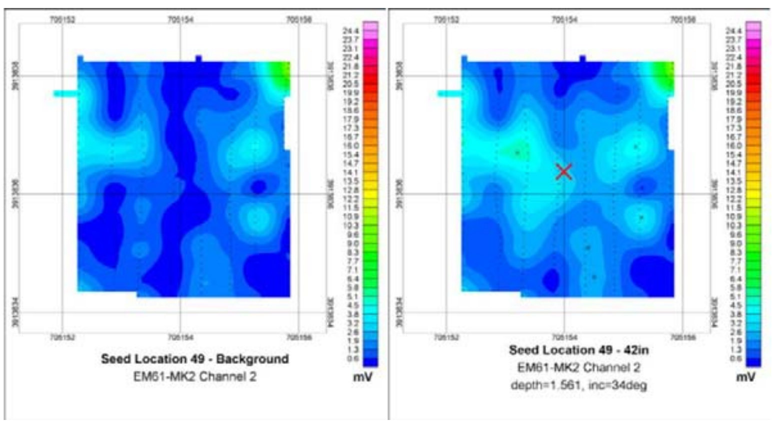

Figure 13. Example of a deep synthetic seed that was not detected due to low signal strength. Left – background signature. Right – background signature with synthetic response of a 4.2in mortar buried 42inches below ground. The red 'X' is the actual seed location.

Actual Target Depth Distributions

The final step in this approach is to evaluate the seeded results within the context of what was actually recovered at the site. In other words, we want to answer the question ‘can we see enough deeper than any of the targets to be satisfied that none were missed?’ A straight forward method of addressing this question is to compare any and all ground truth (i.e., depth distributions as a function of ordnance type) from the site with detection limits from simulations.

Figure 14 present a ground-truth summary of the ordnance depth distributions from Camp San Luis Obispo. For the UXO types included in the study, only one target was recovered at depths of more than 0.5m. The single target was an 81mm mortar found at a depth of 0.52m.

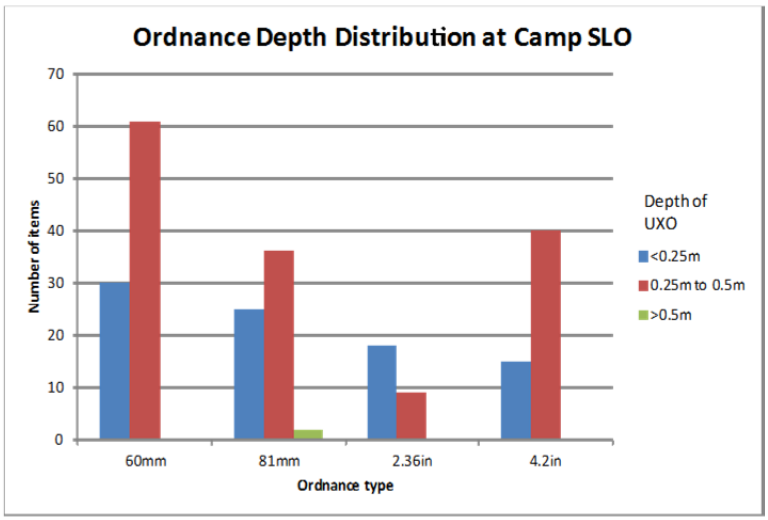

Figure 14. Decision metric distributions from recent ESTCP classification demonstrations. The only target recovered deeper than 0.5m was an 81mm mortar and it was found at 0.52m.

Detection Study Result

Based on this comparison of the actual depth distribution of the UXO recovered at San Luis Obispo and results of the synthetic seed study, it is reasonable to conclude that all of the UXO, at least with regard to the four included in the study, were successfully detected. The deepest target that we know of, based on the available dig results, was an 81mm at a depth of 0.52m. The synthetic seed exercise suggests, however, that 100% of the 81mm mortar seeds were detected for burial depths of up to 0.77m (Figure 9).

Classification Errors

There are two basic types of classification error. Munitions (or other TOI) can be misclassified as clutter. This is referred to as false negative. Or clutter can be misclassified as munitions (false positive). Generally speaking, the consequences are grossly imbalanced. False negatives carry a very high cost associated with leaving a potentially deadly item in the ground, while the cost of a false positive is just digging another hole. Consequently we will focus on the risk of false negatives in the classification process.

Our goal here is to devise procedures for determining the likelihood of false negatives. Initiating events may be inversion failures or unexpected TOI signatures, and we are interested in developing mitigating procedures which can recognize initiating events and procedures for evaluating residual probabilities of misclassified TOI.

The classification process produces a ranked dig list which is based on a decision metric χ determined from target features, where $\chi \to 0$ if there is a good match to a munitions item or other TOI, and $\chi \to 1$ when there is a poor match and the target is likely clutter. Figure 15 shows some decision metric distributions for TOI and clutter items from recent ESTCP classification demonstrations. Classification performance is summarized by the receiver operating characteristic (ROC) curve. Figure 16, from the ESTCP Pole Mountain demonstration, is an example. The ROC is a parametric plot of the percentage of TOI on the site which are correctly classified (Pd) versus the number of incorrectly classified clutter items (Nfa) as the decision metric varies from 0 (marking the beginning of the ROC curve at bottom left point on the plot) to 1 (marking the end of the ROC curve at the top right point of the plot).

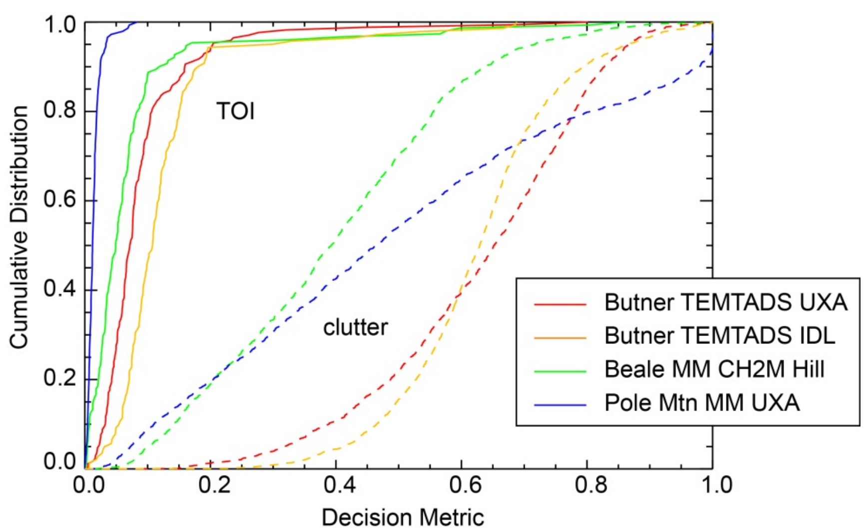

Figure 15. Decision metric distributions from recent ESTCP classification demonstrations.

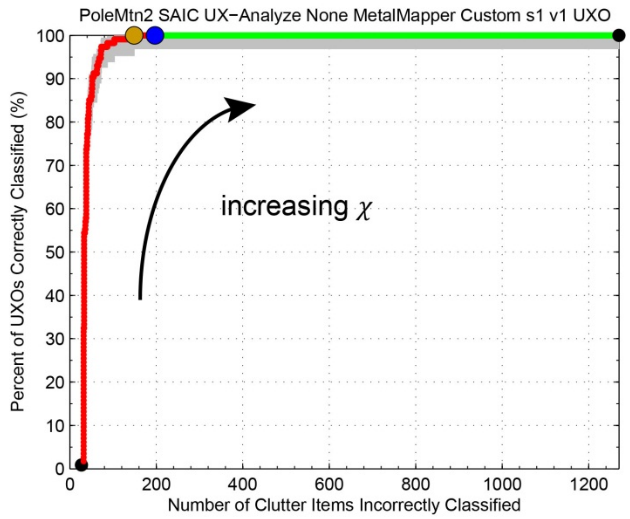

Figure 16. Example ROC curve from Pole Mountain classification demonstration.

The classification process entails setting a stop-dig threshold χ_T on the decision metric. Targets with decision metric $\chi \leq \chi_T$ are classified a possible TOI and will be excavated while those with $\chi > \chi_T$ are classified as likely clutter and will be left in the ground. Consequently the true identities of the targets with metric $\chi \leq \chi_T$ will be known, but the others will not. The residual risk corresponds to the likelihood that any of the targets classified as likely clutter and left in the ground are actually TOI. It depends on the statistics of the decision metric χ for the TOI. We see only samples of the probability distribution of the decision metric for the TOI, $P(\chi|TOI)$. Because there tend to be very few TOI and many clutter items at a typical munitions response (MR) site [2], such samples are generally small.

Probabilistic risk assessment entails

1. Modeling the underlying distributions of test statistics (decision metrics) for TOI and clutter using sample statistics,
2. Calculating the likelihood of misclassified TOI as a function of the amount of clutter that has to be excavated, and
3. Evaluating the costs and/or consequences.

Step 2 can be implemented using extreme value theory or by Monte-Carlo simulations. Here we focus on the application of extreme value theory. The risk is generally defined in terms of the relative costs of false positive and false negative classifications [3]. Here, we will focus on the application of extreme value theory to estimating the likelihood that one or more TOI may be misclassified.

Decision Metrics for Classification

A simple decision metric based on two parameters which was developed in SERDP project MR-2100 [4] turns out to be more useful for our analysis than the decision metric shown in Figure 15. Classification is a matter of deciding whether the object's polarizabilities are TOI-like or clutter-like. Library matching methods employing various procedures to compare polarizabilities of unknown targets with those of targets of interest (TOIs) are commonly used for classification [5, 6]. Our method exploits the fact that an object's polarizability tensor **B**(t) can be represented as a product of two factors: the volume V of the object and a tensor **A**(t) whose eigenvalues $\alpha_j(t)$, i = 1, 2, 3 depend only on the shape and composition of the object. Strictly speaking this is true only for nonmagnetic objects [7] or those with a specific magnetic permeability. However, as a practical matter it appears to be a good representation for typical TOIs and clutter items. Confronted with an unknown target, we then compare its apparent size and EMI "shape" with the sizes and shapes of the TOI.

Given the set (spanning three axes and N time gates) of principal axis polarizabilities β_0 for a TOI and the set of principal axis polarizabilities β for an unknown target, we calculate a size ratio s as

(1)

$$s = median\left(\frac{\sqrt[3]{\beta}}{\sqrt[3]{\beta_0}}\right)$$

where the median is taken over all axes and time gates for which $\beta > 0$. If $\beta < 0$ for more than some threshold fraction (typically 25-50%) of the available terms, then the target is put in the "can't analyze" category. The size mismatch parameter Δ_{size} is then

(2)

$$\Delta_{size} = log(s)$$

which is equal to zero if the EMI sizes of the target and the reference TOI are the same, The shape mismatch parameter Δ_{shape} is determined by comparing the unknown target's polarizability with the reference polarizability scaled by the size mismatch

(3)

$$\Delta_{shape} = \frac{\sum\left|\sqrt[3]{\beta} - s\sqrt[3]{\beta_0}\right|}{\sum\sqrt[3]{\beta}}$$

in which the sums are over all terms with positive β. For each target, size and shape mismatch parameters are calculated for each TOI. Classification is based on thresholding a figure of merit (FOM) parameter or decision statistic

(4)

$$\chi = \min_{TOI}\{|\Delta_{size}| + k\, log(\Delta_{shape})\}.$$

which has been minimized over the set of possible TOI. A parameter value $k \approx 0.3$ seems to give the best classification performance. Because the decision statistic is defined in terms of the logarithm of the shape mismatch $\chi \rightarrow -\infty$ for objects that are a very good match to TOI. We could transform this FOM parameter to a decision metric which runs from 0 to 1 with a transformation based on some cumulative probability distribution model for the FOM. For purposes of analysis, however, it is more convenient to deal with a parameter which can run from $-\infty$ to ∞.

We find similar TOI and clutter distributions of the decision metric defined by equation (4) for the classification demonstrations at Camp San Luis Obispo, Camp Butner and Camp Beale (Figure 17).

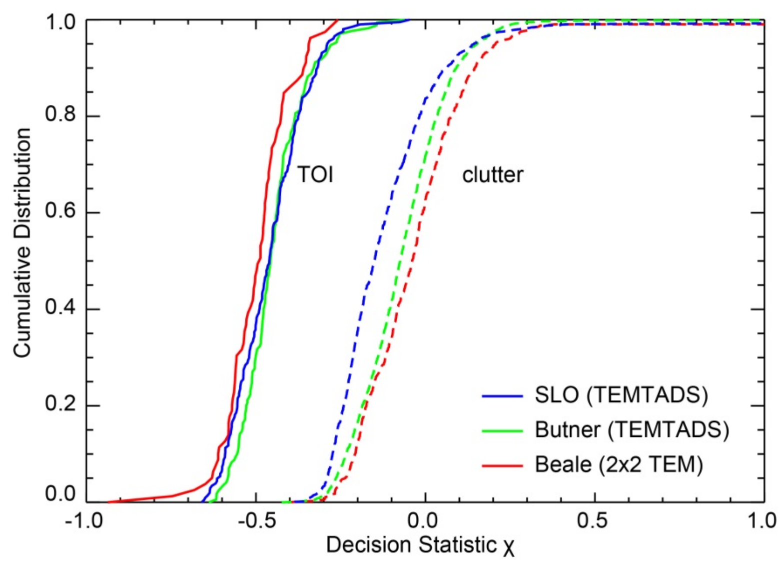

Figure 17. Decision metric distributions for TOI and clutter at Camp San Luis Obispo, Camp Butner and Camp Beale.

Classification performance is controlled by the overlap region between TOI and clutter. The distributions for Camp Beale have the least overlap and correspond to the best classification performance of the three demonstrations.

Extreme Value Distributions

The simple application of statistics can produce inappropriate estimates of the risk that TOI have been misclassified. In a purely statistical approach, the likelihood of TOI misclassification is related to the extreme value statistics for the TOI decision metric (χ) distribution. For example, if there are n TOI, then one measure of the residual risk might be the probability that the maximum value M_n of χ for the TOI items,

(5)

$$M_n = \max(\chi_1, \chi_2, \dots, \chi_n),$$

exceeds the stop-dig threshold χ_T that was selected.

Depending on the tail behavior of the distribution there can be one of three types of extreme value distributions [8, 9]. Figure 18 shows the tails of the TOI distributions from Figure 17. They appear to be bounded by exponential and Gaussian distributions as shown.

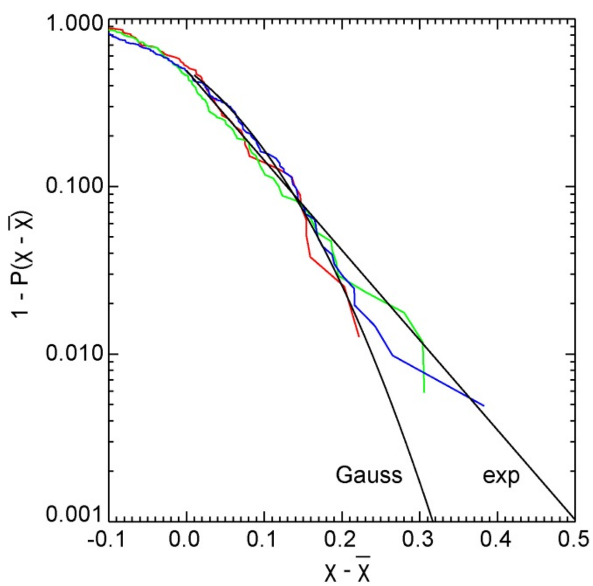

Figure 18. Tails of TOI distributions from Figure 17 compared with Gaussian and exponential distributions.

For exponential-type tails (which includes Gaussian and exponential), with suitable scaling factors a_n (the intensity function) and u_n (the characteristic largest value) the distribution of the maximum approaches the Gumbel distribution

(6)

$$P\{a_n(M_n - u_n) \leq x\} \rightarrow exp(-e^{-x})$$

asymptotically when n is large. If the underlying distribution function for the decision metric χ is Gaussian

(7)

$$F(\chi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\chi} e^{-\frac{1}{2}x^2}\, dx$$

then asymptotically for large n the scaling factors are

(8)

$$a_n = \sqrt{2\, ln(n)} \quad \text{and} \quad u_n = a_n - \frac{1}{2a_n}[ln\, ln(n) + ln(4\pi)].$$

Alternatively, for an exponential distribution

(9)

$$F(\chi) = 1 - e^{-\chi}$$

we have

(10)

$$a_n = 1 \quad \text{and} \quad u_n = ln(n).$$

Figure 19 shows what happens when we apply this to the Camp Beale classification data. The red curves (left scale) show the probability that $M_n > \chi_T$ calculated using the Gaussian and exponential tail distributions shown in Figure 18. The calculations were done for $n = 80$, but the results are not very sensitive to the actual value. The blue curve (right scale) simply shows the number of clutter items with $\chi < \chi_T$. The plot implies that roughly half of the clutter items would be dug at a threshold corresponding to a 1% residual risk of misclassified TOI. We believe this to be misguided. The mathematical formulations are correct but the logic is not. We believe the problem is that the tails in Figure 18 do not reflect random sample variations in the decision metric that should be expected, but rather are due to inversion failures or other pathological events. We consider the question of identifying such failures in the Results and Discussion section below, under the subheading Classification Failures.

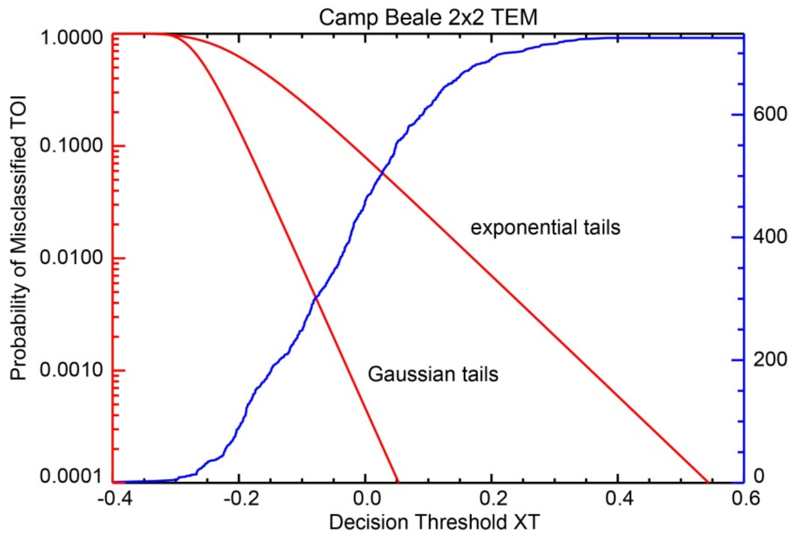

Figure 19. Residual classification risk for Camp Beale as a function of decision threshold based on extreme value statistics (red curves and left scale) and number of clutter digs as a function of threshold (blue curve and right scale).

Results and Discussion

Missed Detections

We presented a practical and straight forward tool for the assessment of missed UXO detections. It involved inserting synthetic signatures, appropriately chosen for the UXO types actually recovered, into the reconnaissance survey data, processing using standard detection schemes and thresholds, and evaluating the results in the light of ground truth information. This approach is attractive because it naturally incorporates site-specific realistic noise levels and sampling schemes.

When the process was applied to ESTCP's Classification Study at San Luis Obispo, CA, it validated the original anomaly selections.

Classification Failures

For our purposes, classification failures are TOI targets that we think ”look” more like what they are not than what they are, and are classified accordingly. In fact, since the decision metric is based on a library of expected TOI signatures, what we really mean is that the target does not look “enough like” a TOI to be considered TOI. But why? Are there unexpected TOI with significantly different polarizabilities than the expected TOI? Are the target's calculated polarizabilities truly a better match to clutter polarizabilities than to polarizabilities for expected TOI at the site? Or are its calculated polarizabilities just very noisy and not a decent match to anything?

The first option (unanticipated TOI) is usually dealt with by using some form of cluster analysis to look for groups of targets with similar signatures. This is pretty straightforward with the size and shape mismatch parameters that go into the decision metric.in equation (4). Figure 20 shows an example of how munitions targets tend to cluster in our size/shape mismatch feature space. We have compared all of the year 1 Pole Mountain MetalMapper targets with one of the seeded 37mm projectiles (target 645). Points are color-coded by target type. The munitions items come in different sizes (only 37mm size targets have size mismatch near zero in this plot) but all have a similar EMI look to them, so an unanticipated munitions item would probably show up with a low shape mismatch to other munitions items. The apparent shape difference between the small ISOs (4 inch pipe sections) and the 37mm projectiles is really a composition difference: The ISOs have relatively thin walls compared to the 37s.

We could quantify the second option (better match to clutter than TOI) by calculating a decision metric based on clutter signatures and comparing it with the decision metric based on the library of expected TOI signatures. However, this would require a clutter library, and some method of evaluating how much signature variability to allow in matching to sample members of the different possible classes of clutter items. These are not currently available, but we can get some

idea how this might play out by using a "clutter library" based on polarizabilities for the targets at the site which were identified as clutter in the post-demonstration excavations. For this we calculate a clutter FOM or decision statistic following equation (4), where now the minimization is over all clutter items (excluding the target whose FOM is being calculated). With χ_{TOI} equal to the TOI-based decision metric and $\chi_{Clutter}$ equal to the clutter-based decision metric, we can define a forced-choice (TOI vs. clutter) decision metric χ_{FC} along a line rotated 45º counter-clockwise from the vertical in the χ_{TOI} - $\chi_{Clutter}$ plane as

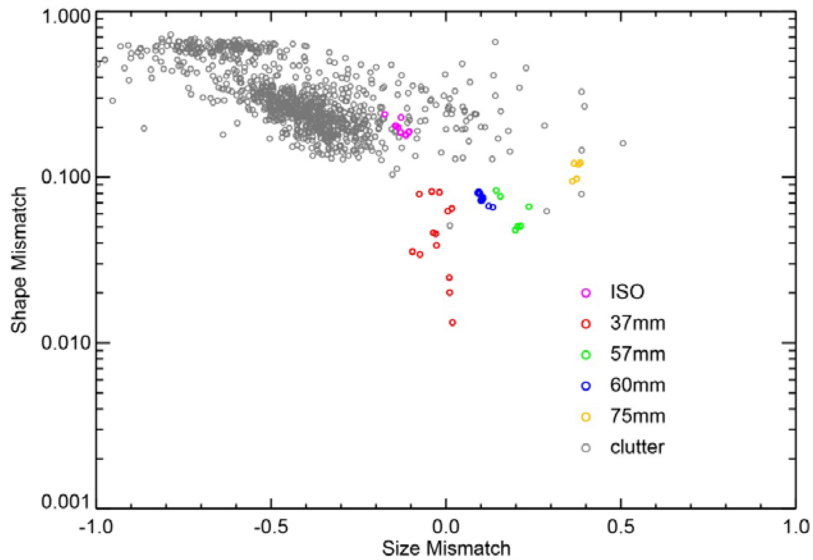

Figure 20. Pole Mountain Year 1 MetalMapper targets compared with one of the seeded 37mm projectiles (target 645) in terms of size and shape mismatch parameters in equation (4).

(11)

$$\chi_{FC} = \chi_{TOI} sin\left(\frac{3\pi}{4}\right) + \chi_{Clutter} cos\left(\frac{3\pi}{4}\right).$$

In this frame, the forced-choice metric $\chi_{FC} \lessgtr 0$ depending on whether the target looks more like the TOI or the clutter.

Figure 21 shows the relationship between χ_{FC} and χ_{TOI} for the Camp Beale man-portable TEM array. The point (χ_{TOI}, χ_{FC}) = (-0.302, -0.090) is relatively far out in the tails of the χ_{TOI} distribution, but on the basis of its χ_{FC} value it is clearly more TOI-like than clutter-like. It represents target BE-903, which is an 81mm mortar. On the other hand the small ISO BE-295, with (χ_{TOI}, χ_{FC}) = (-0.283, 0.008), does not show up as particularly TOI-like or clutter-like. That is because its calculated polarizabilities are very noisy and not a decent match to anything.

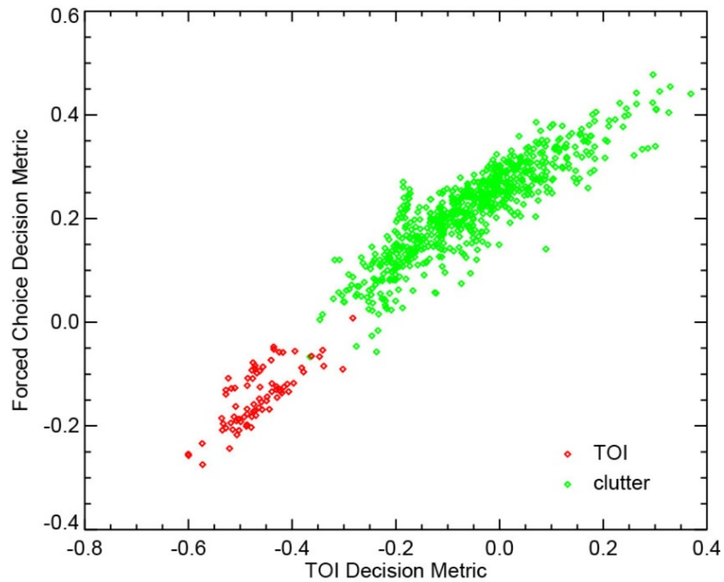

Figure 21. Forced choice (clutter vs. TOI) decision metric compared with TOI-only based decision metric for Camp Beale.

Conclusions and Implications for Future Research

We presented a practical approach to the assessment of residual UXO risk. Specifically, we investigated a method for assessing target detections and devised procedures for determining the likelihood of misclassifying a UXO as clutter. When applied to ancillary sites in the future, these common-sense procedures and associated thought processes are expected to instill confidence with regards to the final decisions and facilitate stakeholder acceptance. If unexpected results are initially revealed, the process should iterate, by changing the detection thresholds or classification logic and digging more targets, until stakeholders, regulators, and the geophysical classification team members agree that all reasonable efforts to mitigate residual UXO risks have been made.

The objective of potential follow-on development and research should be to encode the process and methods into a commercial data processing package and disseminate to the user community for use on ancillary data sets.

Literature Cited

1 ESTCP, “ESTCP Pilot Program Classification Approaches in Munitions Response, San Luis Obispo, California”, " Final Report, May 2010.

2. Roger Young and Arkie Fanning, "Cost Effectiveness Analysis for Cued Interrogation Geophysics at Military Munitions Response Sites," August 11, 2006.

3. Harry L. van Trees, *Detection, Estimation and Modulation Theory*, Part 1, John Wiley and Sons, New York, 1968.

4. Thomas Bell, "Tensor Invariant Processing for Munitions/Clutter Classification," Final Report, SERDP Project MR-2100, July 2013.

5. Leonard R. Pasion, Stephen D. Billings, Douglas W. Oldenburg and Sean E. Walker, "Application of a library based method to time domain electromagnetic data for the identification of unexploded ordnance," Journal of Applied Geophysics, vol. 61 pp. 279–291, 2007.

6. Karl N. Kappler, "An approach to UXO discrimination via polarizability curvematching and feature extraction applied to polarizability," 21st Symposium on the Application of Geophysics to Environmental and Engineering Problems (SAGEEP), Philadelphia, April 2008.

7. L.D. Landau, E.M. Lifshitz and L.P. Pitaevskii, Electrodynamics of Continuous Media, 2nd Edition, §59, Elsevier, Amsterdam, 1984.

8. M.R. Leadbetter, Georg Lindgren and Holger Rootzen, *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, New York, 1983.

9. C.S. Hirtzel, “Analysis of Extreme Values of Natural Processes: Statistical Description of the Maximum,” *Applied Mathematics and Computation*, vol. 15, pp. 283-303, 1984.